

プログラミング必須英単語600+の語彙レベル推測

西野 竜太郎（グローバルゼーションデザイン研究所）

2021-08-23公開

はじめに

プログラミング必須英単語600+は複数のカテゴリーで構成される。そのうち単語の形で掲載されているのは、略語および頭字語を除いた、ベーシック（300語）、アドバンスト（300語）、および前提英単語（100語）の3カテゴリーである。語彙レベルは前提英単語、ベーシック、アドバンストの順に上がるが、学校教育を指標とした場合にどの程度に相当するか（例：高校卒業程度）が不明だった。そこで学校教育での相当レベルを備えた語彙リストと比較することで、大まかではあるが、語彙レベルを推測する。

方法

今回は比較する語彙リストとして、東京外国語大学投野由紀夫研究室で公開している「CEFR-J Wordlist」を利用する。同リストでは語彙がA1～B2のレベル別にまとめられており、「日本の学校教育での相当レベル」も併せて掲載されている。次の4レベルである。

- A1：小学校～中学1年程度
- A2：中学2年～高校1年程度
- B1：高校2年～大学受験レベル
- B2：大学受験～大学教養レベル

手順としては、まずプログラミング必須英単語600+に3つあるカテゴリー（ベーシック、アドバンスト、前提英単語）に入っている各英単語がCEFR-J Wordlistのどのレベルに含まれているかを照らし合わせる。例えば、ベーシックのapplicationはCEFR-J Wordlistでは「B1」、applyは「A2」に含まれている。続いて、どのレベルにいくつ含まれているかをカテゴリーごとに集計する。なお対照時の詳細な方針は付録にまとめている。集計が終わったら含まれる割合に基づいて大まかな語彙レベルを推測する。

結果

集計の結果は表1の通りである。600+の3つのカテゴリーの語彙がCEFR-J Wordlistの各レベル（A1～B2）に含まれる割合および累積割合を示している。「なし」とは語彙リストに掲載されていないという意味である。

| | 前提英単語 | | ベーシック | | アドバンスト | |
|----|-------|------|-------|------|--------|------|
| | 割合 | 累積 | 割合 | 累積 | 割合 | 累積 |
| A1 | 66% | 66% | 10% | 10% | 0% | 0% |
| A2 | 29% | 95% | 28% | 39% | 7% | 7% |
| B1 | 4% | 99% | 28% | 67% | 24% | 31% |
| B2 | 1% | 100% | 13% | 80% | 18% | 49% |
| なし | 0% | 100% | 20% | 100% | 51% | 100% |

表1：3カテゴリーの語彙がCEFR-J Wordlistの各レベルに含まれる割合

次の図1に、表1に記載した割合を積み上げグラフで示す。

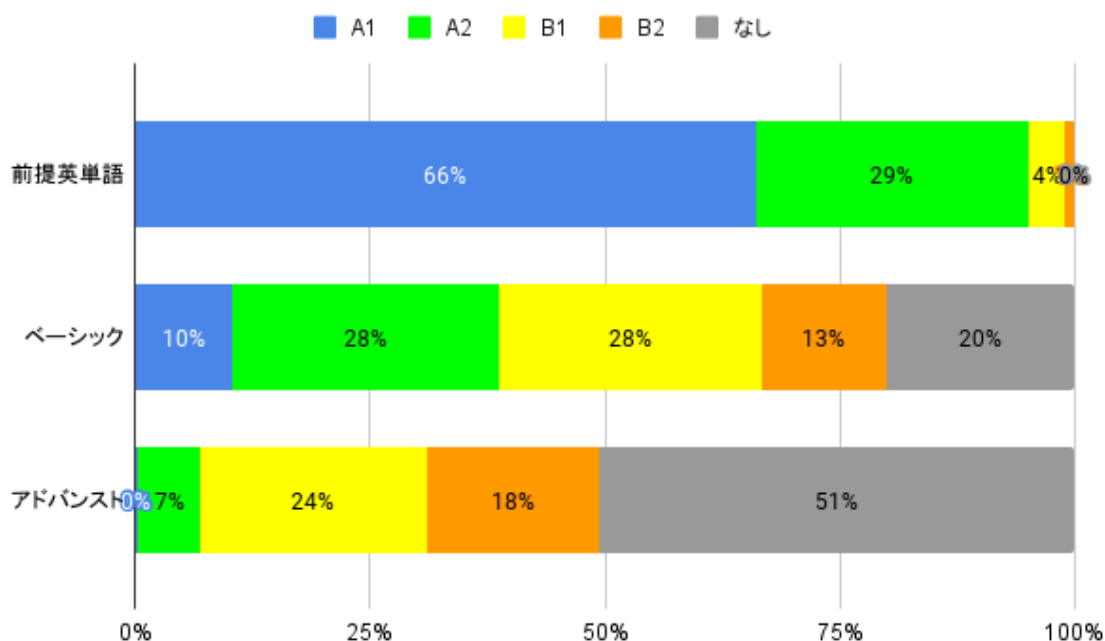


図1：3カテゴリーの語彙がCEFR-J Wordlistの各レベルに含まれる割合

考察

上記の結果に基づき、3つのカテゴリーの語彙レベルを推測してみる。

まず、前提英単語の100語である。CEFR-J WordlistのA1に66%、A2に29%が入っている。A2までの累積で95%である。A2は「中学2年～高校1年程度」とされているため、ほぼ中学校卒業までに習得する単語だと考えられる。そこで前提英単語は「中学卒業レベル」と推測する。

続いて、ベーシックの300語である。累積でB1までで67%、B2までで80%となっている。ベーシックの300語には、CEFR-J Wordlistでは「なし」であるものの、多少プログラミング

経験があれば知っていそうな単語（例：dialog、login）がいくつも含まれている。上記の割合に加えてこの点も考慮すると、高くてもB2の「大学受験～大学教養レベル」程度と考えてもよいかもしれない。そこでベーシックは「高校卒業レベル」と推測する。

最後に、アドバンストの300語である。B2までで累積49%となっている。半分しか入っていないと、さすがにB2レベルと見なすのは難しそうだ。B2の「大学受験～大学教養レベル」からさらなる学習が必要となるだろう。そこでアドバンストは「大学卒業または社会人レベル」と推測する。

注意したい点が1つある。それは、600+で掲載されている意味を学校の英語教育で必ずしも学ぶとは限らない点である。例えば「body」はCEFR-J WordlistでA1レベル、600+でも前提英単語100である。しかしCEFR-J Wordlistでは「体」の意味が想定されているのに対し、600+では「本文、本体」を意味としている。そのため、同じく入門的な英単語であっても、学校教育で習う意味と600+で示す意味との間に差異が生じうる。

おわりに

CEFR-J Wordlistと比較することでプログラミング必須英単語600+の3カテゴリーの語彙レベルを推測した。大まかだが、前提英単語は「中学卒業レベル」、ベーシックは「高校卒業レベル」、アドバンストは「大学卒業または社会人レベル」だと考える。

参考文献

『CEFR-J Wordlist Version 1.6』 東京外国語大学投野由紀夫研究室. (URL: http://www.cefr-j.org/download.html#cefrj_wordlist より2021年7月31日ダウンロード)

付録

対照時の詳細な方針を以下にまとめる。

- 600+の単語に複数の品詞が付与されているケースがある。1つでもCEFR-J Wordlistに対応する場合は含まれるとした。例えばbuildには動詞「構築する」と名詞「ビルド」があるものの、CEFR-J Wordlistには動詞しか含まれていない。
- CEFR-J Wordlistには同じ単語に複数の品詞があり、品詞でレベルが分かれることがある。この場合、低い方を採用した。例えばblockは名詞がA1、動詞がB1であるが、低い方のA1を採用している。
- 600+の単語自体がCEFR-J Wordlistに載っているものの、品詞が合致しない場合は採用しなかった。例えば600+でconsoleは名詞（コンソール）であるが、CEFR-J Wordlistでは動詞でしか掲載がないため、レベルの対応はなしとした。
- 600+の意味とCEFR-J Wordlistが想定している意味は異なることがある。例えば600+でbootは「起動」だが、CEFR-J Wordlistは履物の「ブーツ」が想定されている。今回は意味も含めた対照はせず、表記上で同一であれば対応すると見なした。